Comparison of several fusion paradigms applied to pixel-based image classification

François Leduc
Defence R&D Canada – Valcartier
2459 Boul. Pie-XI Nord
Val-Bélair, QC,
G3J 1X5, Canada
francois.leduc@drdc-rddc.gc.ca

Abstract – This paper is concerned with the development of FuRII, a pixel-based image classification tool developed at DRDC Valcartier. FuRII is based on fuzzy sets and evidence theories and is implemented as an ENVI toolbox. The aim with this tool is to compare several fusion operators and rules in the context of image classification applied to land cover mapping. Several fuzzy fusion operators (conjunctive, disjunctive, adaptive and quantified adaptive fusion) and evidential fusion rules (Dempster, Dubois and Prade, Yager and Smets) are tested. FuRII permits to model imprecise knowledge with membership functions and fusion can be performed directly with membership values or with mass functions. In this later case, a transformation of membership values into basic belief values is computed. Finally, FuRII permits integration of source reliability into the fusion process.

Keywords: image classification, information fusion, fuzzy sets, Dempster-Shafer, evidence, belief, reliability.

1 Introduction

Multisource information fusion is the process of merging several pieces of information in order to obtain the most reliable possible fused picture. Fusion should be a synergetic process, which means that the result should be more accurate than any picture based on an individual source.

To this day, there is no tool commercially available dedicated to information fusion (based of fuzzy sets and evidence theory) applied to land cover mapping or target detection. The development of FuRII (Fuzzy Reasoning applied to Image Intelligence) aims at bridging this gap by making possible testing different fusion schemes. FuRII is experimental and is developed in the IDL programming language and is implemented as an ENVI toolbox.

FuRII is a pixel-based image classification tool that allows knowledge modeling with different types of membership function shapes. Once fuzzy inference is computed, membership values can be fused within the framework of fuzzy sets theory or with dempsterian approaches. In both cases, source reliability can be integrated into the fusion process. If dempsterian fusion is selected, FuRII offers several possibilities for transforming membership values into mass functions.

This paper is arranged as follow. Section 2 gives some theoretical background while section 3 contains a short description of the parameters that can be controlled within FuRII. Section 4 gives a description of the data sets used in this study. Section 5 presents the results obtained with different configurations. Finally section 6 discusses the results and section 7 concludes this document.

2 Theoretical background

2.1 Fuzzy sets

Fuzzy sets theory was proposed by Zadeh in 1965 [1] in order to deal with imprecise information. The fuzzy inference process is the comparison of an observation (a fact) that can be crisp or fuzzy with imprecise information represented by a membership function. The result is a membership value that measures to what extent the fact corresponds to a class according to the feature modeled with the membership function. When considering M features (i.e. spectral bands) and N classes, the fuzzy inference produces a matrix of M by N membership values. In order to decide which class the object belongs to, fusion operators are necessary.

Fuzzy fusion operators include conjunctive, disjunctive, adaptive and quantified adaptive fusion. Conjunctive and disjunctive fusion operators are also referred as triangular norms (t-norms) and triangular conorms (t-conorms) [2][3]. Although there are several types of t-norms and t-conorms, we will limit their definitions with the minimum and maximum operators. Thus the conjunctive fusion can be defined as:

$$T(\mu_A, \mu_B) = \min(\mu_A,\mu_B) \qquad (1)$$

and the disjunctive fusion by:

$$S(\mu_A, \mu_B) = \max((\mu_A,\mu_B) \qquad (2)$$

where μ_A and μ_B are membership values assigned to the same class by two sources S_A and S_B. The adaptive fusion operator was proposed by Dubois and Prade [4] in order to take advantage of both conjunctive and disjunctive operators, while minimizing their negative aspects since conjunctive fusion is considered too severe and disjunctive fusion is considered too permissive. The adaptive fusion (π_{ad}) is expressed as:

$$\pi_{ad} = \max\left(\frac{\pi_{conj}}{h}, \min(1-h, \pi_{disj})\right) \qquad (3)$$

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **JUL 2006**	2. REPORT TYPE	3. DATES COVERED **00-00-2006 to 00-00-2006**

4. TITLE AND SUBTITLE **Comparison of several fusion paradigms applied to pixel-based image classification**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Defence R&D Canada - Valcartier,2459 Boul. Pie-XI Nord,Val-Belair, QC,G3J 1X5, Canada,**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
9th International Conference on Information Fusion, 10-13 July 2006, Florence, Italy. Sponsored by the International Society of Information Fusion (ISIF), Aerospace & Electronic Systems Society (AES), IEEE, ONR, ONR Global, Selex - Sistemi Integrati, Finmeccanica, BAE Systems, TNO, AFOSR's European Office of Aerospace Research and Development, and the NATO Undersea Research Centre. U.S. Government or Federal Rights License

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**8**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

where π_{conj} and π_{disj} are the conjunctive and disjunctive fusion operators and where h is the agreement between sources. Finally, the quantified adaptive fusion can be described as a conjunctive fusion considering the hypotheses that are the most supported [4]. The quantified adaptive fusion (π_{qad}) is given by:

$$\pi_{qad} = \max\left(\frac{\pi_{(n)}^{conj}}{h(n)}, \min(1-h(n), \pi_{(m)}^{conj})\right) \qquad (4)$$

where n is the optimistic evaluation of the number of reliable sources while m is the pessimistic evaluation [5].

2.2 Evidence Theory

Evidence theory was initially proposed by Dempster in 1968 [6] and formalized in 1976 by Shafer [7]. Considering a frame of discernment (Ω) composed of three exhaustive and mutually exclusive hypotheses, H_1, H_2 and H_3, a set ϕ ($\phi = 2^{\Omega}$) called the referential of definition can be composed. This set contains all possible combinations such as:

$\phi = \{(H_1),(H_2),(H_3),(H_1,H_2),(H_1,H_3),(H_2,H_3),(H_1,H_2,H_3),(\varnothing)\}$

where $\varnothing$ represents the hypothesis "other" which is considered in order to respect the exhaustiveness of the propositions. The elements composing the set ϕ are called focal elements and the three elements $\{H_1\}$, $\{H_2\}$ and $\{H_3\}$ are singletons while the others are compound elements. The sum of the masses, calculated over ϕ, must equal one. All elements having a mass (or basic belief) greater the zero make up the body of evidence (N_ϕ). Masses assigned to N_ϕ make up the basic belief distribution or mass function. A mass assigned to the element $\{H_1,H_2\}$ represents the basic belief of being in presence of H_1 or H_2 without being able to discriminate between both elements. A mass equal to unity assigned to element $\{H_1,H_2,H_3\}$ corresponds to total ignorance. Initially, as proposed by Shafer [7], the mass assigned to the empty set is null ($m\{\varnothing\} = 0$) which corresponds to a closed-world paradigm. This means that the solution is necessarily one the initial hypotheses, H_1, H_2 or H_3. This is opposed to the open-world paradigm in which the solution can be something other than the three initial hypotheses. In fact, the dempsterian representation with an open-world context has been formalized by Smets as the Transferable Belief Model (TBM) [8].

The Evidence (or Dempster-Shafer) theory is used to represent uncertain pieces of evidence. Using M sources leads to M mass functions and because fusion is done two- by-two, there are M-1 fusion processes. Among the dempsterian fusion rules there are the Dempster (Ds), the Dubois and Prade (DP), the Yager (Yg) and the Smets (Sm) fusion rules.

The Ds fusion rule [9] is the orthogonal sum ($\oplus$) of two mass functions given by two sources, S_1 and S_2, and uses the conflict as a normalizing factor. In a general way, a mass (m) assigned to a proposition A is given by:

$$m_{S1\oplus S2}(A) = \alpha * \sum_{X\cap Y=A} m_{S1}(X) * m_{S2}(Y) \qquad (5)$$

where α is a constant of normalization given by:

$$\alpha = 1 / (1 - m_\varnothing) \qquad (6)$$

where $m_\varnothing$ is given by the conflicting masses.

With the DP fusion rules [10], conflicting masses are assigned to the propositions implied in the conflict and with the Yg fusion rule [11], conflicting masses contribute to the ignorance by being assigned to the frame of discernment (Ω). Finally, the Sm fusion rule [12] assigns conflicting masses to the hypothesis "other" ($\varnothing$). The Ds, DP and Yg rules belong to the closed-world paradigm while the Sm rules belong to the open-world. Ds and Sm rules are commutative and associative while DP and Yg rules are commutative but not associative.

Once a fusion process is completed, a decision can be based on different criteria such as belief, plausibility and pignistic probability [12],[13].

2.3 From Membership to basic belief values

Figure 1 shows an example of fuzzy inference considering three classes, A, B and C, modeled with three membership functions and an observation X corresponding to a reflectance of 42%. The fuzzy inference produces three membership values (μ) of:

$\mu_A(42) = 0.62$, $\mu_B(42) = 0.49$, $\mu_C(42) = 0.22$.

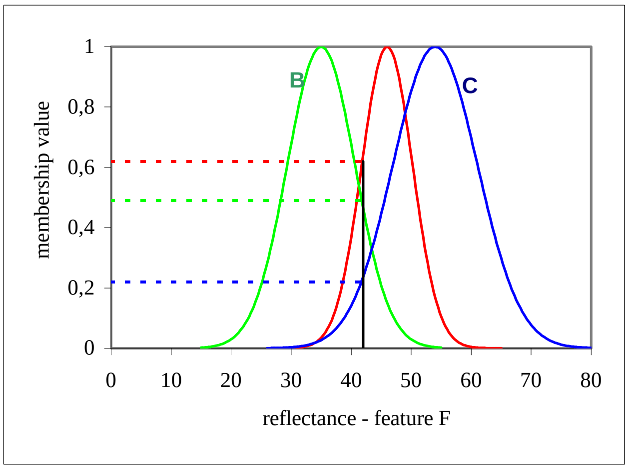

Figure 1 : Illustration of fuzzy inference.

With fuzzy fusion the membership values are used directly but with fusion within the Dempster-Shafer framework a transformation is required. The simplest transformation consists of building Bayesian mass functions where masses are obtained by membership values normalization [14]. A mass to class x is given by:

$$m\{x\} = \mu_x / \sum_{i=1}^{n} \mu_i \qquad (7)$$

where μ is a membership value and n is the number of classes. Thus the three membership values of figure 1 would give the following mass function:

$m\{A\} = 0.47$, $m\{B\} = 0.37$, $m\{C\} = 0.17$.

Transforming membership values into a mass function with such a method gives no advantage if fusing information with the Dempster fusion rule because this rule is conjunctive and would produce similar results as the Zadeh's t-norm. However, this transformation can be advantageous if using the DP or the Yg rules as both can produce compound focal elements. Another advantage of such a transformation resides in the way total ignorance is managed as discussed in section 6.2.

Another approach for transforming membership values into mass functions was proposed in [15],[16] that considers nested focal elements. This method can be illustrated with figure 2. The only singletons of the mass function will be composed of the class having the highest membership value. Other elements will be composed according to their rank after sorting the membership values.

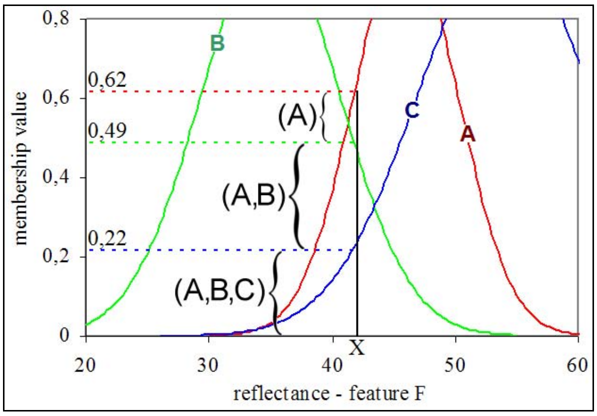

Figure 2 : Nested focal elements derived from membership value. (Enlargement of figure 1).

Thus, the nested mass function will be:

m{A} = 0.62 – 0.49 = 0.13

m{A,B} = 0.49 – 0.22 = 0.27

m{A,B,C} = 0.22 – 0 = 0.22

From here, there are two possibilities to finalize the mass function. In a closed-world context, masses are normalized:

m{A} = 0.13 / 0.62 = 0.21

m{A,B} = 0.27 / 0.62 = 0.44

m{A,B,C} = 0.22 / 0.62 = 0.35

and in and open-world context, the element "other" is added:

m{A} = 0.13

m{A,B} = 0.27

m{A,B,C} = 0.22

m{∅} = 0.38 $(1 - \mu_{max})$

The element ∅ can be seen as a fourth class that is considered only if the highest membership value is lower than 1.

Because focal elements of the mass function are nested, it becomes natural to assign a mass to ∅ that is equal to unity minus the highest membership values $(1 - \mu_{max})$. But with Bayesian mass functions is not as trivial to consider an open-world paradigm as there is no relation between membership values. However, although it can be disputable, we consider the possibility of using an open-world paradigm with Bayesian mass functions by assigning a mass to ∅ that is computed the same way as mentioned above that is:

m{∅} = $1 - \mu_{max} = 1 - 0.62 = 0.38$.

Masses to other elements are computed by normalizing their membership values to μ_{max} so that the sum of masses gives one:

m{A}= 0.29, m{B}= 0.23, m{C}= 0.10, m{∅} = 0.38.

In other words there are four possibilities for transforming membership values into mass functions by selecting between closed and open world contexts and between Bayesian and nested mass functions.

There are also two different ways to consider the open-world paradigm: 1) within the fusion process by using the Sm fusion rule and 2) in the mass functions computation by having the possibility to assign a non-null mass to ∅.

2.4 Sources reliability

If hypotheses are exhaustive, conflict between sources can be explained by one or more sources not being reliable. If there are sources that are not reliable, they should be given less importance in the fusion process. This can be done by using rules such as the trade-off and discount rules [17]. The discount rule can be used on mass functions [18] or on membership values [19]. The discount rule is applied on mass functions by multiplying each evidence value by a reliability coefficient. What has been removed is then added to the whole frame of discernment (Ω):

$$\begin{aligned} m_i^d(A) &= R_i m_i(A), \forall A \subset \Omega \\ m_i^d(\Omega) &= (1 - R_i) + R_i m_i(\Omega) \end{aligned} \qquad (8)$$

where m^d is the discounted mass and where A is any focal element different from Ω. In other words, this rule decreases the importance of evidence values and increases the contribution to "ignorance".

For a source S characterized with a reliability coefficient R_S, membership values are discounted by :

$$\mu_{Ci}' = \max(\mu_{Ci}, 1 - R_S) \qquad (9)$$

where μ_{Ci} is the membership value to the class C_i.

Reliability coefficients, R_i, can be obtained in several ways. One way to compute them is by using a method referred herein as source performance. This method consists of classifying one scene with all spectral bands or features separately. For each feature, a confusion matrix is computed and R_i is given as the overall accuracy. Thus R_i is directly related to the ability of a source to make the good decision. However, this method is influenced by the reliability of the ground truth data.

3 FuRII

FuRII is an experimental tool developed as an ENVI toolbox aiming at testing different fusion configurations. After having loaded imagery and selected samples, the choice of a configuration goes as follow:

1) *Knowledge model*: selection of a membership function shape: Gaussian, Triangular, Trapezoidal or Histogram-shape.

2) *Reliability*: None, Ability to make a decision, Source performance or Class separability.

3) *Fusion operator*: Conjunctive, Disjunctive, Adaptive, Quantified adaptive or Fuzzy evidential fusion.

If the user has selected one of the first four operators, then fuzzy classification begins. If fuzzy evidential fusion is selected, then other parameters need to be selected:

4) *Type of mass functions*: Bayesian or Nested.

5) *World paradigm*: Closed-world or Open-world.

6) *Fusion rules*: Dempster, Dubois and Prade, Yager or Smets. After having selected the fusion rule the classification can begin.

Concerning the reliability, the Ability to make a decision is the difference between the highest membership value (winning class) and the second highest one. Class separability is given by the membership function intersections. In figure 1, the highest intersection is given by membership functions A and C with a value of 0.79. The reliability is then given by 1 – 0.79 = 0.21. Considering membership functions intersections as a measure of confusion has already been presented in [20]. Finally, conjunctive fusion, within FuRII, is implemented as the minimum operator (Zadeh's t-norm) and disjunctive fusion is implemented as the maximum operator (Zadeh's t-conorm).

4 Data sets

In order to analyze the possible configurations within FuRII, four data sets were used. Data set A was composed of a digitized aerial color photography (three bands) of an airport context. Spatial resolution is 64 cm.
Data sets B and C are concerned with a forested environment located in Saskatchewan, Canada. Data set B is composed of the six Landsat 5 Thematic Mapper multispectral bands while data set C is composed of the three bands TM3, TM5, TM7 plus three texture features (based on co-occurrence matrices). Spatial resolution is 30 m. Texture was computed on the first component resulting from a principal components analysis calculated with the six Landsat bands. The three texture features (contrast, variance and entropy) were computed with a 7x7 kernel, directional invariant with a distance of 1 and a 32 gray level quantization.
Data set D is composed of a digital aerial photo composed of four multispectral bands (blue, green red and near infrared) concerning a parking lot containing civilian and military vehicles. Spatial resolution is 15 cm. For this data set, in order to reduce inner class variability, a morphological dilation filter was applied to the four bands. Table 1 presents the description of the data sets and Figure 3 shows previews of parts of these data sets.
For all data sets the knowledge (membership functions) was computed from samples drawn on imagery. Validation of classification was done with the use of thematic maps. In the case of data sets A and D, the ground truth was obtained by manually digitizing the objects. In the case of data sets B and C the ground truth was obtained by combining a maximum likelihood classification with a thematic map produced in the framework of the BOREAS project [21]. The considered ground truth was composed of the correctly classified pixels.

5 Results

As can be seen from section 3, there are many possible configurations with FuRII. But preliminary results allowed drawing preliminary conclusions that help in guiding the rest of the analysis. The first observations are the followings:

- Among the fuzzy fusion operators, quantified adaptive fusion performs better than the three other operators;
- Fuzzy fusion is much faster than evidential fusion;
- Using Bayesian mass functions (MF) gives very similar results then nested MF when using Ds, DP or Yg rules;

Table 1: Description of the data sets.

Data set A	Data B
Airport	Forest
3 bands: Blue, Green, Red	6 bands: TM1, TM2 TM3, TM4, TM5, TM7
426 x 421 pixels Resolution: 64 cm	1152 x 883 pixels Resolution: 30 m
3 classes: Aircraft, Tarmac, Grass	4 classes: Conifers, Mixed, Deciduous, Water
Data set C	Data set D
Forest	Parking lot
6 bands: TM3, TM4, TM5, contrast, variance, entropy	4 bands: Blue, Green, Red, NIR
1152 x 883 pixels Resolution: 30 m	354 x 263 pixels Resolution: 15 cm
4 classes: Conifers, Mixed, Deciduous, Water	3 classes : Cars, Military vehicles, Asphalt

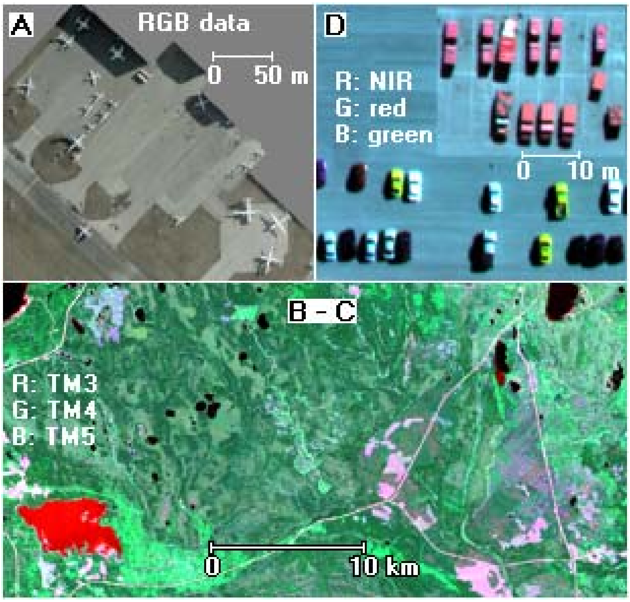

Figure 3 : Previews of the data sets used in this study.

- The SM rule performs better with nested MF;
- Open-world MF give poorer performances than closed-world MF;
- The Sm fusion rules gives the poorest performance. Its performance is even "catastrophic" when combined with open-world MF;
- DP and Yg fusion rules give very similar performances. Because these rules are not associative, they perform better if sources are fused in order of increasing reliability;

- If sources are fused in such an order, usually the DP and Yg rules perform better than the Ds rule;
- When integrating reliability in the fusion process, the Ds, DP and Yg rules give better performance than if reliability is not used .

According to these observations, the following discussion will present results for these fusion schemes: conjunctive (Con), disjunctive (Disj), adaptive (Ad) and quantified adaptive (Qad) operators and Dempster (Ds), Dubois and Prade (DP) and Yager (Yg) rules. For these three dempsterian rules, nested MF built in a closed-world context were used. Membership functions are histogram-shaped and reliability is evaluated as source performance which is computed from samples. Finally, because DP and Yg fusion operators are not associative, bands were sorted and fused by order of increasing reliability.
Classification with FuRII requires samples to be collected in order to compute knowledge (statistics) and membership functions for each class and each band. The samples are also used as test sites for computing reliability coefficients with the sources performance method. These coefficients are presented in Table 2.

Table 2: Reliability coefficients of the bands/features composing the four data sets.

Data set A		Data set B		Data set C		Data set D	
R	.842	TM1	.536	TM3	.568	R	.626
G	.875	TM2	.588	TM4	.989	G	.502
B	.908	TM3	.568	TM5	.899	B	.644
		TM4	.989	cont.	.685	NIR	.908
		TM5	.899	var.	.684		
		TM7	.790	entr.	.699		

Although there are several ways to present classification results such as confusion matrices, kappa coefficients, percentage of unclassified pixels, errors of omission and of commission, we limit the results here to overall accuracies (OA) in order to be concise and succinct.
Concerning data set A, the best performance was obtained with the Qad fusion (Table 3). No dempsterian fusion rules perform better even when integrating sources reliability. Conjunctive and disjunctive fusion produced very similar results. The main difference between both operators is that the pixels that are unclassified with the conjunctive operator became confused with the disjunctive fusion. An unclassified pixel is a pixel for which each class is characterized by at least one null membership value while a confused pixel is assigned to more than one class with the same membership value.

Table 3: Data set A. Overall accuracies for several fusion operators.

Con	Disj	Ad	Qad	Ds	DP	Yg
0.627	0.623	0.697	0.751	0.736	0.742	0.742
Using reliability (sources performance)						
0.719	----	----	----	0.746	0.748	0.747

When integrating reliability into the fusion process, the four fusion rules (Con, Ds, DP, Yg) see there performance increase but none of them performed better than Qad fusion. This may be explained by the low number of sources (3 bands) leading to only two fusion steps and by the high values of the reliability coefficients.
With the data set B (Table 4), the Qad fusion was the best of the fuzzy operators but its performance was exceeded by the DP and Yg rules. Moreover, when reliability is integrated into the fusion process even the conjunctive and the Ds rules performed better. This is in relation with the high level of concordance (Table 10) between sources.

Table 4: Data set B. Overall accuracies for several fusion operators.

Con	Disj	Ad	Qad	Ds	DP	Yg
0.741	0.213	0.780	0.808	0.771	0.847	0.891
Using reliability (sources performance)						
0.894	----	----	----	0.898	0.914	0.915

Data set C is the one that is composed with the more heterogeneous features that are highly uncorrelated. This is due to the nature of texture measurements that are implied in the decrease of sources' concordance. This is illustrated by the poor performance of the conjunctive operator and the Ds rule (Table 5). The DP and Yg rules are less sensible to this lack of concordance because conflicting masses are assigned to compound focal elements so that almost no class is eliminated during the fusion process. When reliability is used, then performances of the conjunctive and Ds rule become much better. These results will be discussed in section 6.3.

Table 5: Data set C. Overall accuracies for several fusion operators.

Con	Disj	Ad	Qad	Ds	DP	Yg
0.283	0.265	0.439	0.681	0.417	0.826	0.863
Using reliability (sources performance)						
0.894	----	----	----	0.877	0.897	0.903

Finally with data set D there are some significant differences in the results. First of all, the integration of reliability in the fusion process gives poorer performance than no reliability integration (Table 6). This might be explained by the fact that for this data set, three of the four reliability coefficients are close to 0.5 (Table 2) leaving an important part of the decision process to high uncertainty (see section 6.4). Also with this data set, the Smets fusion was tested and it is the rule that gave the poorest performance.

Table 6: Data set D. Overall accuracies for several fusion operators.

Con	Disj	Ad	Qad	Ds	DP	Yg	Sm
.806	.480	.853	.860	.836	.814	.841	.410
Using reliability (sources performance)							
.768	----	----	----	.831	.821	.828	.668

For this data set, other fusion configurations were tested. First, Bayesian mass functions were tested (Table 7) and results are similar to those obtained with nested mass

functions (Table 6) except for the Sm rule which performs better when using nested mass functions.

Table 7: Data set D. Overall accuracies for Dempsterian fusion rules using Bayesian mass functions.

Ds	DP	Yg	Sm
0.836	0.821	0.824	.012
Using reliability (sources performance)			
0.826	0.823	0.823	.027

The open-world paradigm was also tested with the Dempster and Smets fusion rules (Table 8). Combining open-world mass functions with the Sm rule is really not adequate. It seems also better to consider open-world mass functions with the Dempster fusion rule then considering closed-world mass functions with the Smets fusion rule.

Table 8: Data set D. Results obtained with other fusion configurations.

Nested mass functions			
	Ds/ow	Ds/ow$^{(R)}$	Sm/ow
O.A.	.698	.631	.075
∅	28%	29%	93%
Bayesian mass functions			
	Ds/ow	Ds/ow$^{(R)}$	Sm/ow
O.A.	.698	.640	0
∅	28%	29%	100%

$^{(R)}$: reliability used in fusion, ow: open-world, ∅: percentage of pixels classified as "other"

6 Discussion

6.1 Synergy

The first question that might be of interest is: are the fusion processes synergetic? In order to be so, a fusion process must produce an overall accuracy greater than the best reliability coefficient composing a data set because these coefficients are obtained by computing overall accuracies of each individual band. But comparing fusion OA with reliability coefficients is not so straightforward here because fusion OA is computed with ground truth data and reliability coefficients are computed from samples. So in order to compare the synergetic aspect of fusion, reliability coefficients were computed from the same ground truth data as the one used for computing fusion OA. These new coefficients are presented in Table 9. These values can now be compared to results presented in section 5.

Table 9: Overall accuracies for each band/feature computed with ground truth data.

Data set A		Data set B		Data set C		Data set D	
R	.390	TM1	.413	TM3	.319	R	.318
G	.638	TM2	.494	TM4	.896	G	.317
B	.714	TM3	.319	TM5	.678	B	.476
		TM4	.896	cont.	.132	NIR	.755
		TM5	.678	var.	.128		
		TM7	.567	entr.	.203		

With the data set A (Table 3), only the conjunctive, disjunctive and adaptive operators are not synergetic because the OA obtained with these operators is lower than 0.714 (Table 9).

With data set B (Table 4), only Ds, DP and Yg rules integrating reliability produced OA greater than 0.896. Other fusion operators are not synergetic.

With data set C (Table 5) only DP and Yg rules are synergetic while with data set D (Table 6) only the disjunctive fusion and the Smets rule are not.

These results do not permit to draw a conclusion about which fusion operator or rule is best in terms of synergy. The difference in the results may be explained by the number of sources, the agreement between them and by their reliability coefficients.

Table 10 shows statistics about sources agreement computed for each data set with all processed pixels. Note that agreement is computed the same way as fuzzy conjunctive fusion (Zadeh's t-norm) as this operator corresponds to the maximum of consensus reached by sources. Table 10 shows the mean and median agreement value and the value at the 75th and 90th percentiles. Although the agreement may help in interpreting the results, it is important to note that it is not because sources agree strongly that they give the good decision.

Table 10: Statistics of source agreement for the four data sets.

	A	B	C	D
Median	.220	.486	---	.349
75 perc.	.529	.624	.137	.518
90 perc.	.714	.757	.353	.667
Mean	.301	.408	.099	.345
Std.dev.	.288	.287	.187	.239

6.2 Total ignorance

There is one important point to mention about the Dempster fusion rule and its conjunctive behaviour [22]. If, for example, four classes are being analyzed with three sources and that each class is being characterized by at least one null membership value, a pixel will remain unclassified because no consensus is reached resulting in a null agreement. Moreover, if one source gives all null membership values, again it means that no consensus can be reached. The pixel then remains unclassified. But lets look at things differently with the membership values of Table 11 where source S_3 gives all null values. With conjunctive fusion a pixel characterized with such membership values would remain unclassified. Same thing with the Ds rule because no consensus can be reached and also because no mass function can be built with source S_3 (equation 7). In this example, the source S_3 can not participate in the classes discrimination thus it contributes to ignorance. In that sense, with source S_3 all the mass can be given to total ignorance by assigning it to the element {A,B,C,D} which corresponds to the whole frame of discernment (Ω). This way, source S_3 has no effect on the fusion process and the Ds rule may be used with the sources S_1 and S_2. So while with this example, the fuzzy conjunctive fusion would result in an unclassified pixel, the Ds rule can classify the pixel using sources S_1 and S_2. This way of processing sources contributing to total ignorance is implemented within

FuRII and it explains why the conjunctive fusion and the Ds rule do not produce same results.

Table 11: Example of mass functions obtained from membership values assigned to four classes according to three sources.

	A	B	C	D
	Membership values			
S_1:	0	0.4	0.3	0.8
S_2:	0.9	0.7	0.5	0.2
S_3:	0	0	0	0
	Corresponding mass functions			
S_1:	m{B} = .27 m{C} = .20 m{D} = .53			
S_2:	m{A} = .39 m{B} = .30 m{C} = .22 m{D} = .09			
S_3:	m{Ω} = 1			

6.3 Impact of reliability coefficients

Concerning the reliability coefficients, one might raise some questions about their meanings. According to the coefficients of Table 2 the texture features would be better than spectral bands TM1, TM2 and TM3 for forest classes discrimination. But recall that these coefficients are computed from samples selected by a user and that human beings tend to select homogeneous samples that may not reflect the true complexity of classes. When coefficients are computed from more objective data (i.e. ground truth data), their relative values change in a significant way (Table 9). We can see that the texture features (data set C) become very unreliable and that their impact on classification results should be almost null. These more realistic reliability values explain the lower performance of the conjunctive and Ds operators.

The question now is what would be the classification results if reliability coefficients of Table 9 were used in the fusion process ? This was tested with data sets C and D for some of the fusion rules. Concerning data set C (Table 12), the use of the new coefficients almost eliminates the effect of the three texture features which makes the results more dependent on the three spectral bands. Moreover, band TM3 has a relatively low coefficient which gives more importance to bands TM4 and TM5. This is reflected in the result obtained with the conjunctive fusion where result of is identical to the reliability of the band TM4. It demonstrates that most of the decision is based on this band.

Concerning data set D, the use of reliability coefficients of Table 9 with the dempsterian fusion rules improves the results (Table 13). Actually it modifies the conclusion from which data set D was the only one where the use of reliability did not improve the results. This example shows that the value of these coefficients directly influences the results.

Table 12: Data set C. Overall accuracies obtained with the reliability coefficients of Table 9.

Con	Ds	DP	Yg
0.896	0.913	0.913	0.913

Table 13: Data set D. Overall accuracies obtained with the reliability coefficients of Table 9.

Con	Ds	DP	Yg
0.752	0.847	0.838	0.846

6.4 Comments

The integration of reliability coefficients into the fusion process may improve the classification accuracies but, in fact, it improves the results only if some sources are reliable enough. The quality of the results will depend on the coefficients values. Table 14 shows an example of a mass function that is adjusted according to three reliability coefficients values. If the reliability is high (0.9), the adjustment brings almost no difference. If the reliability is low (0.1), the source has almost no impact on the fusion because almost all the belief is placed on Ω. Finally, when reliability is average (0.5), belief values are significantly decreased and an important part of belief is placed on Ω. While it contributes partly to ignorance it may still participate in making the wrong decision.

Table 14: Example of a weighted mass function according to three reliability coefficients values.

Mass function		Reliability coefficients		
class	masses	R	R	R
A : B: C:	0.6 0.3 0.1	0.1	0.5	0.9
Adjusted mass functions				
	A:	0.06	0.3	0.54
	B :	0.03	0.15	0.27
	C :	0.01	0.05	0.09
	Ω :	0.9	0.5	0.1

This example shows that it is better two have one unreliable source than several mid-reliable sources.

6.5 Computing time

Table 15 shows a comparison of computing times for different fusion processes concerning data sets A and B. We can see that fuzzy fusion is much faster then evidential fusion. Fuzzy fusion is linearly dependant on the number of pixels while the evidential fusion is dependant on the number of pixels and on the number of bands. This relation is explained by the fact that the evidential fusion is done two-by-two so if using N sources, the number of fusion operations is N-1 for each pixel. Moreover, once fusion is done, a pignistic probability is computed in order to make the final decision. So the processing load for each pixel is much heavier with evidential fusion.

Table 15: Computing time in minutes for the fusion processes of data sets A and B.

	Data set A		Data set B	
	nr	SP	nr	SP
Con	19s*	57s*	2	8
Qad	24s*	---	3	---
Ds	3	10	55	182
DP	3	11	70	185
Yg	3	11	65	215

SP: sources performance, nr: no reliability

**: duration in seconds*

7 Conclusion

In this study many fusion configurations were tested with an experimental pixel-based image classification tool named FuRII. Knowledge about objects of interest is modeled with membership functions. Classification by fusion can be done directly with membership values obtained by fuzzy inference. Fusion can also be done in a dempsterian framework which requires a transformation of membership values into mass functions. Sources reliability can also be integrated into the fusion process.

The first conclusion to be drawn is that Dempsterian fusion rules are much slower than fuzzy fusion operators.

If the question of processing time is important, quantified adaptive fusion can be used with relatively good confidence. If a bit more time is available fuzzy conjunctive fusion is appropriate if reliability is used.

If even more time is available, the Yager fusion rule integrating reliability may be used safely.

However, the results of the fusion process will depend on the reliability coefficients values. If all sources are completely unreliable, it becomes impossible to make a decision. If reliability coefficients are close to 0.5, classification results becomes highly uncertain. So if reliability coefficients are low, it may be necessary to question the pertinence of the data used.

References

[1] L.A. Zadeh, *Fuzzy Sets*, Information and Control, 8:338-353, 1965.

[2] D. Dubois and H. Prade, *Combination of fuzzy information in the framework of possibility theory*, Data Fusion in Robotics and Machine Intelligence, Abidi and Gonzalez eds., Academic Press, 481-505, 1992

[3] L. Gacôgne, *Éléments de logique floue*, Hermès, 1997.

[4] D. Dubois and H. Prade, *Possibility theory and data fusion in poorly informed environments*, Control Engineering Practice, 2(5):811-823, 1994.

[5] D. Dubois and H. Prade, *La fusion d'informations imprécises*, Traitement du Signal, 11(6):447-458, 1994.

[6] A.P. Dempster, *A Generalization of Bayesian Inference*, Journal of the Royal Statistical Society, Series B, 30:205-247, 1968.

[7] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, 1976.

[8] P. Smets, *The Transferable Belief Model*, Artificial Intelligence, 66:191-234, 1996.

[9] L.A. Zadeh, A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination, The AI magazine, summer 1986, pp. 85-90.

[10] D. Dubois, and H. Prade, *Evidence, knowledge and belief functions*, International Journal of Approximate Reasoning, 6(3):295-319, 1992.

[11] R.R. Yager, *On the Dempster-Shafer framework and new combination rules*, Information Sciences, 41:93-137, 1987.

[12] P. Smets, *The combination of evidence in the transferable belief model*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(5):447-458, 1990.

[13] P. Smets, *Decision Making in the TBM: the Necessity of the Pignistic Transformation*, International Journal on Approximate Reasoning, 38:133-147, 2005.

[14] L. Roux, Fusion d'informations multisources pour la classification d'images satellite, Ph.D. thesis, Université Paul Sabatier, France, 1997.

[15] F. Leduc, Prototype de système expert pour l'analyse de l'imagerie satellitale dans un contexte de mise à jour des cartes de l'inventaire forestier, Ph.D. thesis, Université de Montréal, 2003.

[16] F. Leduc, B. Solaiman and F. Cavayas, Combination of fuzzy sets and Dempster-Shafer theories in forest map updating using multispectral data, SPIE AeroSense 2001.

[17] G.L. Rogova and V. Nimier, Reliability in Information Fusion: literature survey, Fusion 2004.

[18] E. Lefevre, O. Colot and P. Vannoorenberghe, *Belief function combination and conflict management*, Information fusion, 3:149-162, 2002.

[19] D. Dubois and H. Prade, Possibility theory – an approach to the computerized processing of uncertainty, Plenum Press, 1998.

[20] F. Leduc, *Feature space optimization prior to fuzzy image classification*, Fusion 2004, June 28 – July 1, Stockholm, Sweden.

[21] J. Cihlar, F.G. Hall, M. Apps, B. Goodison, H. Margolis, A. Nelson and P.J. Sellers, *Remote Sensing in BOREAS*, Proceedings of 16th Canadian Remote Sensing Symposium, 1993.

[22] I. Bloch, *Information Combination Operators for Data Fusion: A Comparative Review with Classification*, IEEE Transactions on Systems Man and Cybernetics – Part A: Systems and humans, 26(1):52-67, 1996.